



Analyse bayésienne de courbes de croissance de poulets par un modèle hiérarchique défini à partir d'une équation différentielle stochastique à processus d'erreur autocorrélée

Bénédicte Fontez Nguyen The, Dolores Batonon, Fabien Campillo, Philippe
Lescoat

► To cite this version:

Bénédicte Fontez Nguyen The, Dolores Batonon, Fabien Campillo, Philippe Lescoat. Analyse bayésienne de courbes de croissance de poulets par un modèle hiérarchique défini à partir d'une équation différentielle stochastique à processus d'erreur autocorrélée. 44. Journées de la Société Française de Statistique (SFdS 2012), May 2012, Bruxelles, Belgique. 6 p. hal-01190283

HAL Id: hal-01190283

<https://hal.science/hal-01190283>

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANALYSE BAYÉSIENNE DE COURBES DE CROISSANCE DE POULETS PAR UN MODÈLE HIÉRARCHIQUE DÉFINI À PARTIR D'UNE ÉQUATION DIFFÉRENTIELLE STOCHASTIQUE À PROCESSUS D'ERREUR AUTOCORRÉLÉE

Bénédicte Fontez ^{1,2} & Dolores I. Batonon ^{2,3} & Fabien Campillo ² & Philippe Lescoat ³

¹ *Supagro Montpellier, UMR Mistea, F-34060 Montpellier, France,
Benedicte.Fontez@supagro.inra.fr*

² *INRA, UMR729 Mistea, F-34060 Montpellier, France*

³ *INRA, UR83 Recherches avicoles, F-37380 Nouzilly, France,
Dolores.Batonon@tours.inra.fr; Philippe.Lescoat@tours.inra.fr*

⁴ *MODEMIC Project - INRA/INRIA, UMR Mistea, F-34060 Montpellier, France,
Fabien.Campillo@inria.fr*

Résumé. Les courbes de croissance animale peuvent souvent être définies à partir d'Equations Différentielles Ordinaires(EDO) avec récemment l'utilisation de modèles mixtes comme dans Strathe (2009), *et al.* (2010). Or, la croissance est un processus aléatoire et Donnet *et al.* (2010) ont proposé l'utilisation d'Equations Différentielles Stochastiques (EDS) pour définir le terme de régression d'un modèle hierarchique bayésien. Le processus d'erreur de cette EDS est un mouvement brownien avec accroissements indépendants. Nous proposons d'utiliser le processus d'Ornstein-Uhlenbeck à la place pour modéliser un processus d'erreur autocorrélée plus conforme à la réalité biologique. Les deux modèles ont été implémentés et comparés pour modéliser la croissance de poulets dans un modèle hiérarchique bayésien. Ils ont aussi été comparés au modèle plus simple utilisant les équations différentielles ordinaires. De plus, hormis l'intérêt de modéliser les courbes individuelles, nous avons étudié l'avantage de ces modèles pour comparer des conditions d'élevage et prédire le poids à l'abattage.

Mots-clés. Courbes de croissance, modèle hiérarchique bayésien, équations différentielles stochastiques, processus d'erreur autocorrélée.

Abstract. Animal growth curves may be well described by differential equations which can be used in a mixte model approach, see Strathe (2009), *et al.* (2010). As growth is a stochastic process, Donnet *et al.* (2010) suggested a bayesian hierarchial model where the regression term arised from a stochastic differentiel equation. These corresponding stochastic growth curves were defined with an uncorrelated error process. Instead, we suggested using an Orstein-Uhlenbeck process to model an autocorrelated process which seemed biologically more sensible. The two models were computed and compared on broiler chicken growth data. Eventually, we studied if those models were useful to compare experimental designs and also to predict final average body size.

Keywords. Growth curves, hierarchical Bayes model, stochastic differential equations, autocorrelated error process.

1 Introduction

Les courbes de croissance des poulets sont souvent définies à partir d'Equations Différentielles Ordinaires (EDO) avec récemment l'utilisation de modèles mixtes par Strathe (2009), *et al.* (2010). Or, le poids d'un animal varie selon différents processus faisant intervenir par exemple le génotype, l'alimentation, les pratiques d'élevage et les conditions climatiques. Ainsi, la croissance d'un poulet est un processus aléatoire car, représentant l'évolution du poids d'un individu donné en fonction du temps. Donnet *et al.* (2010) ont proposé l'utilisation d'Equations Différentielles Stochastiques (EDS) pour définir le terme de régression d'un modèle mixte. Le processus d'erreur de cette EDS est un mouvement brownien avec accroissements indépendants. Nous proposons d'utiliser le processus d'Ornstein-Uhlenbeck à la place pour modéliser un processus d'erreur autocorrélée qui semble plus conforme à la réalité biologique. Les deux modèles ont été implémentés pour l'équation de Gompertz dans un modèle hiérarchique bayésien. Ces deux modèles ont aussi été comparés au modèle plus simple utilisant les équations différentielles ordinaires. De plus, hormis l'intérêt de modéliser les courbes individuelles, nous avons étudié l'avantage de ces modèles pour prédire le poids à l'abattage par rapport au modèle naïf (utilisant uniquement le poids à l'abattage sans tenir compte des mesures antérieures).

2 Présentation des données

Face aux problèmes de bien-être soulevés par l'intensification de la filière avicole, un arrêté ministériel transpose en France la directive européenne 2007/43/CE relative à la protection des poulets destinés à la consommation de viandes. Dès lors, les éleveurs sont dans l'obligation de déclarer à l'avance la densité de poulets maximale qu'ils vont adopter et le poids à l'abattage de leurs animaux. Le plan d'expérience de cette étude a été défini pour un lot d'animaux placés dans des conditions où la densité au sol et l'hétérogénéité initiale du poids varient. Les poids de 1440 poulets standard ont été mesurés à 5 âges différents : $t = 0, 9, 21, 35, 41$ jours après éclosion (cf. Figure 1). Ils ont été répartis dans 48 parquets selon un plan factoriel (2×2). Le facteur densité d'élevage (20 ou 40 poulets/parquet) et le facteur hétérogénéité du coefficient de variation du poids initial (6% ou 12% selon les parquets) définissent une variable traitement à 4 modalités. Le plan d'expérience est équilibré car pour chaque modalité, 12 parquets ont été constitués.

3 Présentation des modèles

Le modèle de Gompertz est usuellement utilisé pour la croissance du poulet d'après Anthony *et al.* (1991). Notons Z_t le poids d'un poulet au temps t . (Z_t) est un processus aléatoire défini à partir de l'équation différentielle du modèle de Gompertz. Trois versions ont été considérées :

1. EDS avec fonction de volatilité affine et processus de Wiener d'après Donnet *et al.* (2010)

$$dZ_t = B C e^{-Ct} Z_t dt + \gamma Z_t dW_t, \quad Z_0 = A e^{-B}, \quad (1)$$

où W_t est un processus de Wiener standard (ou mouvement brownien), $W_t \sim N(0, t)$.

2. EDO si on considère que le terme $\gamma = 0$.
3. EDS avec fonction de volatilité affine et processus d'Ornstein-Uhlenbeck

$$dZ_t = B C e^{-Ct} Z_t dt + \gamma Z_t d\xi_t, \quad Z_0 = A e^{-B}, \quad (2)$$

$$d\xi_t = -\alpha \xi_t dt + dW_t, \quad \xi_0 \sim N(0, \frac{1}{2\alpha}) \quad (3)$$

où ξ_t est un processus d'Ornstein-Uhlenbeck. Ce processus aussi appelé reverting mean process (retour vers la moyenne) permet de modéliser un processus d'erreur autocorrélée.

Les trois équations différentielles admettent une solution explicite qui se calcule pour les équations différentielles stochastiques à partir des formules d'Ito présentées dans Øksendal (1995) ou Kloeden et Platen (1992), en posant $X_t = \ln(Z_t)$. Chaque poulet i élevé dans un parquet k avec le traitement j a une trajectoire de croissance en poids, aléatoire, notée (Z_{ijkt}) . A partir de la formule d'Ito, on a une solution explicite pour la trajectoire de chaque poulet en échelle logarithme (X_{ijkt}) .

Une version discrétisée pour les cinq âges observés ($t_0 = 0, t_1 = 9, t_2 = 21, t_3 = 35, t_4 = 41$) s'écrit explicitement comme :

$$\begin{aligned} (\ln(y_{ijkt_0}), \dots, \ln(y_{ijkt_4}))' &= (X_{ijkt_0}, \dots, X_{ijkt_4})' + \varepsilon_{ijk} \\ (X_{ijkt_0}, \dots, X_{ijkt_4})' &= (\ln A_{ijk} - B_{ijk}, \dots, \ln A_{ijk} - B_{ijk} e^{-C_{ijk} t_4})' \\ &\quad - \frac{1}{2} \gamma^2 (t_0, t_1, \dots, t_4)' + \eta_{ijk} \end{aligned}$$

où $(\ln(y_{ijkt_0}), \dots, \ln(y_{ijkt_4}))$, (ε_{ijk}) et $(\ln A_{ijk}, B_{ijk}, C_{ijk})$ désignent respectivement pour un poulet, le vecteur des poids observés (en échelle log), le vecteur des erreurs résiduelles et le vecteur des paramètres du modèle de Gompertz. Le fait d'élever les poulets par parquet introduit une hiérarchie et une corrélation possible entre les valeurs des poulets d'un même parquet au même âge, soit $cor(\varepsilon_{ijkt}, \varepsilon_{i'jkt}) = \rho$. Les poulets d'un parquet n'étant pas supposés indépendants, nous raisonnerons par la suite sur le vecteur des poulets par parquet. Précisons qu'un parquet est identifié par la combinaison du traitement j ($j=1$

à 4) et du numéro du parquet k ($k=1$ à 12). Les aléas ε_{ijk} et η_{ijk} sont indépendants et de loi normale. La principale différence entre les deux EDS réside dans la définition de la matrice de variance-covariance ($\gamma^2 T_{ijk}$) de η_{ijk} :

EDS (1)	EDS (2)
$T_{ijk} = \min(t_l, t'_l)_{0 \leq l, l' \leq 4}$	$T_{ijk} = \frac{1}{2\alpha} e^{-\alpha t_l - t_{l'} _{0 \leq l, l' \leq 4}}$

Dans l'équation (1), plus un animal prend de l'âge et plus la variabilité dans sa courbe de croissance individuelle augmente. Il paraît plus cohérent, comme dans l'équation (2) que la variance reste constante ($\gamma^2/(2\alpha)$). De plus, dans l'équation (1), le terme à $t = 0$ est supposé fixe et les poids ultérieurs sont indépendants de ce terme (variance et covariance nulles) . Or le poids à la naissance est l'aboutissement d'un processus de croissance (dans l'oeuf) et devrait plutôt être considéré comme aléatoire. De même, biologiquement, les futurs poids devraient être corrélés à cette valeur initiale. Dans l'équation (2), $X_0 = \ln A - B + \xi_0$ est aléatoire et la corrélation entre les poids décroît exponentiellement en fonction du temps qui les sépare, en partant de $t = 0$, ce qui semble plus cohérent sur un plan biologique.

A partir de cette écriture discrétisée, on peut définir un modèle hiérarchique bayésien :

$$\begin{aligned}
\ln(y_{jk}) \mid \sigma^2, \rho, X_{jk}, \phi_{jk} &\sim N((X'_{jk})', \sigma^2 \Sigma(\rho)) \\
X_{jk} \mid \phi_{jk}, \gamma^2, \alpha, &\sim N(f(\phi_{jk}, \gamma), \gamma^2 T(\alpha)) \\
\phi_{jk} \mid \mu_j, \Omega &\sim N((\mu'_j, \dots, \mu'_j)', I_n \otimes \Omega) \\
\mu_j &\sim N((m_a, m_b, m_c)', v)
\end{aligned}$$

avec les notations $\phi_{ijk} = (\ln A_{ijk}, B_{ijk}, \ln C_{ijk})'$, $\phi_{jk} = (\phi'_{1jk}, \dots, \phi'_{njk})$ et $X_{jk} = (X_{1jkt_0}, \dots, X_{1jkt_4}, \dots, X_{njk t_0}, \dots, X_{njk t_4})'$. Le paramètre α dans le modèle ci-dessus correspond au paramètre d'autocorrélation du processus d'Ornstein-Uhlenbeck, il n'intervient que pour l'équation différentielle stochastique (2). Dans le cas de l'EDO, le modèle se simplifie car le terme γ est supposé nul et les X_{jk} ne sont plus aléatoires, la trajectoire de la croissance est fixe selon une équation de Gompertz. Pour les termes de variance-covariance ($(\sigma^2, \rho, \gamma^2, \alpha$ et $\Omega)$), les lois *a priori* usuelles inverse gamma et inverse Wishart ont été utilisées. L'estimation des modèles a été réalisée à partir d'un échantillonneur de Gibbs. Certains termes ont nécessité d'implémenter un algorithme de Métropolis Hasting dans l'échantillonneur de Gibbs.

4 Méthodes de validation

Les modèles ont été comparés à partir des distributions prédictives *a posteriori* pour le poids moyen par parquet à l'âge d'abattage situé à $t_4 = 41$ jours (posterior predictive

p-value). Leurs résidus définis selon Carlin et Louis (2000) et la matrice de variance-covariance *a posteriori* des η_{ijk} ont été explorés ainsi que la précision de l'estimation du poids moyen par parquet à l'âge d'abattage. Ces précisions ont été comparées à celle du modèle naïf ($y_{jkt_4} \sim N(moy_{jk}, \sigma^2 \Sigma_0)$) et $moy_{jk} \sim N(moy_j, \tau_j^2)$ avec $moy_j \sim N(moy, \tau)$. Enfin, l'impact du traitement sur la courbe de croissance a été étudié.

Références

- [1] Anthony, N. B., Emmerson, D. A., Nestor, K. E., Bacon, W. L., Siegel, P. B. and Dunnington, E. A. (1991), Comparison of growth curves of weight selected populations of turkeys, quail and chickens, *Poult. Sci.*, 70, 13–19.
- [2] Donnet, S. Foulley, J-L and Samson, A. (2010), Bayesian analysis of growth curves using mixed models defined by stochastic differential equations, *Biometrics*, 66, 733–741.
- [3] Kloeden, P. E. and Platen, E. (1992), *Numerical Solution of Stochastic Differential Equations*, Springer.
- [4] Øksendal, B. (2000), *Stochastic Differential Equations - An Introduction with Applications*, Springer.
- [5] Strathe, A. B. (2009), *Stochastic modelling of feed intake, growth and body composition in pigs*, PhD Thesis, Faculty of life sciences, University of Copenhagen.
- [6] Strathe, A. B., Danfaer A., Sorensen H. and Kebreab, E. (2010), A multilevel nonlinear mixed effects approach to model growth in pigs, *J. Anim. Sci.*, 88, 638–649.

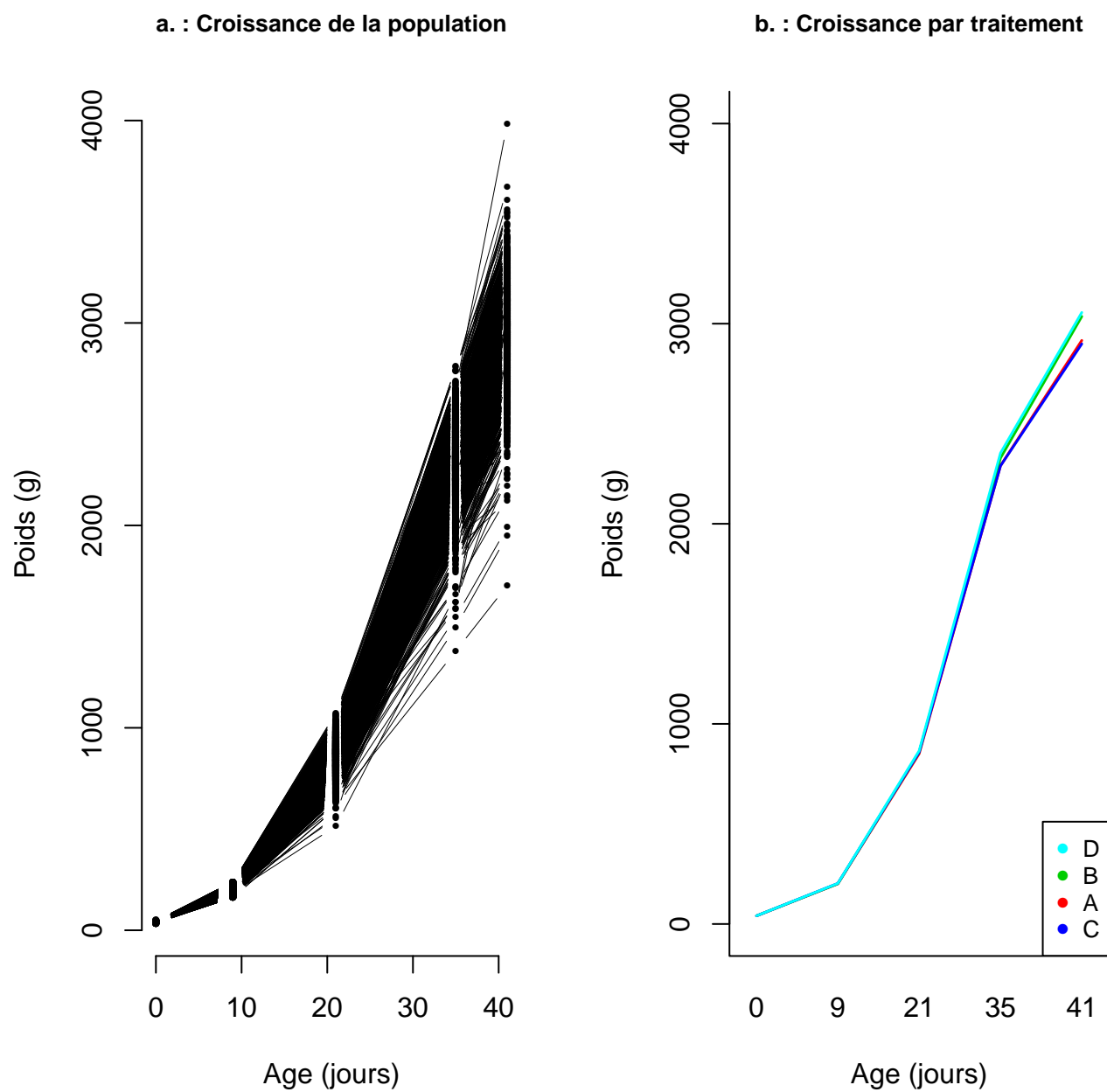


FIGURE 1 – Présentation des données